Polygenic Scores Associated With Educational Attainment in Adults Predict Educational Achievement and ADHD Symptoms in Children

Eveline L. de Zeeuw,1,2* Catharina E.M. van Beijsterveldt,1,2 Tina J. Glasner,1 M. Bartels,1,2 Erik A. Ehli,3 Gareth E. Davies,3 James J. Hudziak,4 Social Science Genetic Association Consortium, Cornelius A. Rietveld,5 Maria M. Groen-Blokhuys,1,2 Jouke Jan Hottenga,1 Eco J.C. de Geus,1,2 and Dorret I. Boomsma1,2

1Department of Biological Psychology, VU University, Amsterdam, the Netherlands
2EMGO+ Institute for Health and Care Research, VU University Medical Centre, Amsterdam, the Netherlands
3Avera Institute for Human Genetics, Avera McKennan Hospital & University Health Center, Sioux Falls, SD
4University of Vermont, College of Medicine, Burlington, VT
5Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands

Manuscript Received: 15 March 2014; Manuscript Accepted: 2 June 2014

The American Psychiatric Association estimates that 3 to 7 per cent of all school aged children are diagnosed with attention deficit hyperactivity disorder (ADHD). Even after correcting for general cognitive ability, numerous studies report a negative association between ADHD and educational achievement. With polygenic scores we examined whether genetic variants that have a positive influence on educational attainment have a protective effect against ADHD. The effect sizes from a large GWA meta-analysis of educational attainment in adults were used to calculate polygenic scores in an independent sample of 12-year-old children from the Netherlands Twin Register. Linear mixed models showed that the polygenic scores significantly predicted educational achievement, school performance, ADHD symptoms and attention problems in children. These results confirm the genetic overlap between ADHD and educational achievement, indicating that one way to gain insight into genetic variants responsible for variation in ADHD is to include data on educational achievement, which are available at a larger scale.

© 2014 Wiley Periodicals, Inc.

Key words: polygenic scores; genetics; ADHD; educational achievement; school performance

INTRODUCTION

The American Psychiatric Association (APA) estimates that 3 to 7 per cent of all school aged children are diagnosed with attention deficit hyperactivity disorder (ADHD) [APA, 2000]. These children have difficulties with inattention, impulsivity and/or hyperactivity interfering with their daily functioning. Children with ADHD have, on average, lower general cognitive ability than controls [Frazier et al., 2004]. At school, they are more likely to repeat a grade and are more often referred to specialized education [Biederman How to Cite this Article:


Conflict of Interest: The authors declare that there is no conflict of interest

Grant sponsor: Organization for Scientific Research (NWO); Grant number: 056-31-001; Grant sponsor: Spinozapremie; Grant number: NWO/SP/1 56-464-14192; Grant sponsor: “Twin-family database for behavior genetics and genomics studies”; Grant number: NWO 480-04-004; Grant sponsor: European Research Council; Grant number: ERC-230374; Grant sponsor: Genetic influences on stability and change in psychopathology from childhood to young adulthood; Grant number: NWO/ZonMW 91210020; Grant sponsor: Genomics of Developmental Trajectories in Twins; Grant number: ARRA RC2 2MH08995.

*Correspondence to: E.L. de Zeeuw, MSc, Department of Biological Psychology, VU University, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands E-mail: el.de.zeeuw@vu.nl

Article first published online in Wiley Online Library (wileyonlinelibrary.com): 5 July 2014 DOI 10.1002/ajmg.b.32254

© 2014 Wiley Periodicals, Inc.
et al., 1996]. In addition, lower educational attainment is negatively related to numerous outcomes, including earnings [Julian and Kominski, 2011] and health [Mackenbach et al., 1997]. The phenotypic association between ADHD and general cognitive ability and ADHD and educational achievement is estimated to be around \(-0.30\) [Kuntsi et al., 2004; Polderman et al., 2010]. This correlation is also shown longitudinally, attention problems assessed at an earlier age predicted lower general cognitive ability and educational achievement years later [Polderman et al., 2006, 2010]. Twin and family studies showed that a substantial part of the (longitudinal) associations seems to be due to shared genetic influences [Kuntsi et al., 2004; Saudino and Plomin, 2007; Polderman et al., 2006]. However, there are no studies that examined at the genotype level whether ADHD and educational achievement share the same common genetic variants.

ADHD in children is approximately 75 per cent heritable [Faraone et al., 2005] and several candidate genes have been identified [Neale et al., 2010a, 2010b; Mick et al., 2010; Stergiakouli et al., 2012], but the largest meta-analysis of genome-wide association (GWA) studies with 2,064 trios, 896 cases and 2,455 controls has not led to the discovery of causal genetic variants associated with ADHD [Neale et al., 2008]. One of the explanations is that ADHD is a highly complex disorder caused by many common genetic variants with small effects. The non-significant single nucleotide polymorphisms (SNPs) probably captured relevant genetic variation, but sample sizes have not been large enough to detect these small effects [Neale et al., 2010b; Flint and Munafò, 2013]. The same is true for educational achievement. It is a trait which is heritable, with heritability estimates of 60–70 per cent in children in current Western society [Bartels et al., 2002; Kovas et al., 2007], with a substantial phenotypic and genetic correlation with general cognitive ability [Bartels et al., 2002], and approximately 40 per cent in adults [Rietveld et al., 2013]. A large GWA study of educational attainment (126,559 adult individuals) revealed genome-wide significant associated genetic variants with a largest estimated effect of 0.02 per cent [Rietveld et al., 2013]. In an additional analysis, Rietveld et al. [2013] combined the effect of all genetic variants, including non-significant variants, and explained approximately 20 per cent of the variance in educational attainment, indicating that educational attainment too is a very polygenic phenotype.

By using polygenic scores the information from non-significant genetic variants can be used to test whether these genetic variants with small effects may actually explain a part of the variance [International Schizophrenia Consortium, 2009]. Polygenic scores also allow for exploration of the underlying etiology of the association between two phenotypes, such as, ADHD and educational achievement. Here, it is expected that genetic variants associated with one phenotype, will explain part of the variance in the other phenotype. Recently, this method was applied by Lencz et al. [2014], who compared polygenic scores, consisting of genetic variants related to general cognitive ability, between schizophrenic patients and controls, showing that the schizophrenic group had lower polygenic scores than the control group. This suggests that some of the genetic variants are involved in both general cognitive ability and schizophrenia. No study has yet used polygenic scores to determine whether the same common genetic variants underlie both ADHD and educational achievement.

The expression of the genotype of an individual may depend on age, with different genes influencing a phenotype in adults and children. General cognitive ability becomes more heritable over time although the influencing genes remain relatively stable [Haworth et al., 2010; Deary et al., 2012; Franic प, et al., 2014]. In childhood, there is a large overlap between genes that have an influence on educational achievement at the start of primary school and genes that are influencing the trait at the end of primary school [Kovas et al., 2007]. However, less is known about the genetic stability of educational achievement from childhood into adulthood. Therefore, we first determined whether the same genetic variants contribute to educational attainment in adults and educational achievement and school performance in children. In the present study, polygenic scores, based on the effect sizes from the meta-analysis of educational attainment by Rietveld et al. [2013], were calculated for children of primary school age and used to explore the association between educational achievement and ADHD symptoms and attention problems.

**METHODS AND MATERIALS**

**Participants**

The Netherlands Twin Register (NTR), established around 1987 by the Department of Biological Psychology at the VU University Amsterdam, registers approximately 40 per cent of all multiple births in the Netherlands [Boomsma et al., 2006; van Beijsterveldt et al., 2013]. The parents of the twins receive a survey about the development of their children every two years until the twins are 12 years old. Since 1999, at age 7, 9 and 12, when the twins attend primary school, parents are asked for their consent for the NTR to approach the teacher(s) of their children with a survey. Genotypes and data for educational achievement, school performance, ADHD symptoms and/or attention problems at age 12 were available for 2,133 children. Data were excluded if a child had a non-European ancestry (N = 106) or had a disease or handicap that interfered severely with daily functioning (N = 38). The ancestry was determined on the basis of a child’s genotype data by using the 1000 Genomes dataset as a reference and was confirmed by the data on birth country of the parents of the child. Eight principal components of the 1000 Genomes cluster the European populations together and a child was labeled non-Dutch ancestry when the child was an outlier on one of these principal components [Abdellaoui et al., 2013]. This resulted in a total sample of 1,989 12-year-old children with genotype data. The children belonged to 1,030 families.

**Measurements**

Educational achievement was assessed by a score on a national standardized test of educational achievement, which is administered in the last grade of primary school at approximately 80 per cent of all schools in the Netherlands [CITO, 2002]. The test consists of multiple choice items in four different subjects, namely Arithmetic, Language, Study Skills and Science and Social Studies. The first three subscales are combined into a Total Score, which is standardized on a scale between 500 and 550. Due to the fact that the Total Score of the educational achievement test has also been
collected in other surveys send to the parents, teacher and the children themselves, there are more data available for the Total Score than for the subscales. School performance was assessed by teacher ratings of three educational domains, namely Arithmetic, Language and Reading, with two versions of the teacher survey. In the first version (birth cohorts 1989–1993), teachers could choose up to six educational domains and rate the proficiency of the students on a five-point scale from 1 (insufficient) to 5 ((very) good). In the second version (birth cohorts 1994–2000), teachers rated the proficiency of the students in four predefined educational domains on the same five-point scale. Due to the free choice in the first version of the survey, the sample size for the teacher ratings differs across educational domains.

Attention Problems (AP) were assessed, by teachers, with the Teacher Report Form (TRF) and, by mothers, with the Child Behavioral Check List (CBCL) [Achenbach, 1991]. The TRF AP scale consists of 20 items rated on a 3 point scale from 0 (not at all or never) to 2 (very true or often). The CBCL AP scale consists of 10 items, which overlap with the TRFAP scale. ADHD symptoms were assessed, by teachers, with the ADHD index of the short version of the Conners’ Teacher Rating Scales - Revised (CTRS-R) and, by mothers, with the ADHD index of the short version of the Conners’ Parent Rating Scales - Revised (CPRS-R) [Conners et al., 1998]. Both scales consist of 12 items, of which 6 overlap, rated on a 4 point scale from 0 (not true or never) to 3 (completely true or very often). Sum scores were computed when subjects had no or a limited number of missing items on a scale (no more than two for ADHD symptoms and mother-rated AP and no more than three for teacher-rated AP). A missing item on a scale was imputed by taking the rounded average of the scale for that child.

**Genotype Data**

**Discovery sample.** The GWA meta-analysis on educational attainment conducted in a discovery sample of 126,559 individuals is described in Rietveld et al. [2013]. Educational attainment was analyzed as a quantitative measure defined as an individual’s total years of schooling (EduYears) and as a binary measure defined as whether or not an individual had a college degree (College). The GWA meta-analysis discovery sample included adult NTR participants, who could be related to the children in the sample used in the present study. Therefore, the meta-analysis was carried out again and effect sizes were estimated based on a sample without the NTR individuals.

**Target sample.** Selection of NTR samples for genotyping gave preference to samples which were collected in unselected groups of children who participated in cognition studies [Polderman et al., 2006; Hoekstra et al., 2007], in MRI projects [van Soelen et al., 2012], and in a study of blood group chimerism [van Dijk et al., 1996]. Approximately one third of the children took part in a study [Derks et al., 2008] that selected children for the presence or absence of high AP/ADHD symptom scores. Children with longitudinal DNA samples, or whose parents could be genotyped, also were included [Scheet et al., 2012]. The remaining samples were selected based on the availability of longitudinal phenotype data collected in NTR survey studies. There were no significant differences for educational achievement, school performance, attention problems and ADHD symptoms between the genotyped group (N = 1989) and the group of children without genotype data (N = 16581).

NTR individuals were genotyped on Illumina and Affymetrix 6.0 platforms. Data were phased using Mach 1.0 and genotype imputation was performed on a larger sample with Minimac using the 1000 Genome Phase 1 Integrated Release Version 3 Haplotypes (b37, HG19). For the quality control (QC) of the single nucleotide polymorphisms (SNP), criteria were a Hardy-Weinberg equilibrium (HWE) P-value > 0.00001, minor allele frequency (MAF) > 0.01, call rate > 0.95, Mendel error rate < 0.02 and allele frequency difference with reference < 0.20 and, for C/G and A/T, SNPs were only included with a MAF < 0.35. For the QC of the samples, criteria were a missing rate < 0.10, heterozygosity −0.10 < F < 0.10, consistency between reported gender and sex chromosome genotypes, consistency of expected and observed family relatedness between samples and Mendelian error rate < 0.02.

**Polygenic Scores**

The SNPs were clumped using the linkage disequilibrium (LD) based result clumping procedure in Plink [Purcell et al., 2007]. All SNPs were included with the default settings of a LD threshold based on a R^2 of 0.50 and a distance threshold of 250 kb to ensure that only nearly independent SNPs were included in the calculation of the polygenic scores. Only SNPs overlapping between the discovery and the target sample were included in the clumping procedure. The remaining criteria for the SNPs were a MAF > 0.01, an info score > 0.40 in both the discovery and the target sample, a difference in the frequency of the reported minor alleles < 0.15, and, for C/G and A/T SNPs, a MAF < 0.35. The clumping procedure led to a total number of 343,388 and 352,397 SNPs for the calculation of the polygenic scores based on, respectively, EduYears and College.

The polygenic scores were then calculated in the Plink program [Purcell et al., 2007] for each individual by multiplying the number of observed effect alleles with the effect size found in the meta-analysis and summed over all SNPs [International Schizophrenia Consortium, 2009]. The effect size of a SNP was calculated by multiplying the METAL [Willer et al., 2010] z-statistic with the square root of twice the MAF times the major allele frequency [Rietveld et al., 2013]. For each individual, two polygenic scores were calculated; one based on the effect sizes for EduYears and the other on those for College. The polygenic scores were calculated for different subsets of SNPs, selected on the bases of their p-value in the discovery sample, with thresholds of 5 × 10^{-8}, 1 × 10^{-5}, 0.001, 0.01, 0.05, 0.1, 0.5 and 1.0.

**Statistical Analyses**

Educational achievement, school performance, AP and ADHD symptoms were regressed on the EduYears and College polygenic scores in linear mixed models using the Statistical Package for the Social Sciences (SPSS 20) [IBM, 2011]. For each analysis, the predictor and the outcome measure were standardized within each subset of children with data available on both. To correct for dependency of the observations due to family clustering an additive genetic variance component was included as a random effect based on the family pedigree and dependent on zygosity.
Covariates included in the analyses were gender, three principal components, reflecting ancestry differences within the Dutch population, six principal components, reflecting ancestry differences based on the 1000 Genomes project, one principal component, correcting for a possible batch effect and a number of dummy variables to indicate the specific subset in which individuals were genotyped, to correct for differences in genotyping quality across sets [Abdellaoui et al., 2013]. The amount of variance (R²) explained by the polygenic scores was calculated by squaring the standardized regression coefficient. The total number of independent dimensions in the outcome data was extracted from the correlation matrix of the phenotypes with the MatSpD program developed by Nyholt [2004]. The phenotypes contained 7 independent dimensions [Li and Ji, 2005] and therefore a P-value of 0.007 (0.05/7) was considered significant.

RESULTS

First, it was tested whether the polygenic scores based on educational attainment, as obtained in an adult sample, predicted educational achievement and school performance in children. Both genotype and phenotype data were available for educational achievement in Arithmetic (N=745, Mean = 59.3, SD = 27.2), Language (N=746, Mean = 58.7, SD = 27.2), Study Skills (N=744, Mean = 61.0, SD = 26.4), Science and Social Studies (N=662, Mean = 57.5, SD = 27.8) and the Total Score (N=1365, Mean = 537.7, SD = 8.4) and for school performance in Arithmetic (N=901, Mean = 3.83, SD = 1.13), Language (N=847, Mean = 3.84, SD = 0.97) and Reading (N=470, Mean = 3.79, SD = 1.08). The standardized regression coefficients are reported in Table I. The results are given for the polygenic scores based on years of education (EduYears) and for completion of College (College). The polygenic scores significantly predicted educational achievement and school performance at certain thresholds. All significant effects were in the expected direction. Polygenic scores based on EduYears and College, significantly predicted the score of the ADHD index at certain p-value thresholds (Table I). All significant effects were in the expected direction. Higher polygenic scores were associated with a lower score on the ADHD index, especially for the larger sample of mother ratings. For AP, the effects were less clear. There was only one polygenic score, based on College at the threshold of suggestive genome-wide significance, for AP rated by teachers. The highest proportion of variance explained by the EduYears (Fig. 1) and College (Fig. 2) polygenic scores was at different thresholds (pT) for teacher-rated AP (EduYears: R² = 0.006, P = 0.022 at pT = 0.1; College: R² = 0.014, P = 5 × 10⁻⁴ at pT = 1 × 10⁻⁵) and ADHD index (EduYears: R² = 0.011, P = 0.016 at pT = 0.1; College: R² = 0.021, P = 0.002 at pT = 0.001) and for mother-rated AP (EduYears: R² = 0.002, P = 0.098 at pT = 1.0; College: R² = 0.005, P = 0.010 at pT = 1 × 10⁻⁵) and ADHD index (EduYears: R² = 0.014, P = 2 × 10⁻⁴ at pT = 0.1; College: R² = 0.009, P = 0.003 at pT = 1 × 10⁻⁵).

To investigate whether the non-normality of the school performance and ADHD phenotypes influenced the results, all analyses for school performance, attention problems and ADHD symptoms were also conducted after normalizing the data in PRELIS [Joreskog and Söbom, 2002]. PRELIS transforms the data by fitting an inverse normal density function to the ranked data. This normalization did not alter the results of the polygenic score analyses. Almost exactly the same pattern of significant regression coefficients was observed (results are available upon request from the first author).

DISCUSSION

Up to ~3 per cent of the variance in educational achievement and up to ~4 per cent of school performance in children was explained by the polygenic scores that were based on educational attainment in adults. The polygenic scores predicted, in general, educational achievement in children to the same extent as educational attainment in adults [Rietveld et al., 2013]. This finding is consistent with numerous results from the genetics literature on general cognitive ability, in which the involvement of the same genetic variants is found in childhood and adulthood [Hoekstra et al., 2007; Haworth et al., 2010; Franic et al., 2014]. The current study suggests that this is also true for educational achievement, as polygenic scores based on a general measure of educational attainment obtained in adults predicted children’s performance across a number of educational domains.

This is the first study that demonstrated genetic associations between ADHD and educational achievement using polygenic scores. The polygenic scores, representing educational attainment in adults, explained up to ~2 per cent of the variance in attention problems and ADHD symptoms in 12-year-olds, indicating that some of the genetic variants that have a positive influence on educational attainment have a protective effect against ADHD. This confirms, at the measured genotype level, the existence of shared genetic effects accounting for the negative association between ADHD and educational achievement, previously found in twin research [Saudino and Plomin, 2007; Polderman et al., 2013].
<table>
<thead>
<tr>
<th></th>
<th>Arithmetic</th>
<th>Language</th>
<th>Study skills</th>
<th>Science and social studies</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>p-value</strong></td>
<td><strong>β</strong></td>
<td><strong>p</strong></td>
<td><strong>β</strong></td>
<td><strong>p</strong></td>
<td><strong>β</strong></td>
</tr>
<tr>
<td><strong>Threshold</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
</tr>
<tr>
<td>0.01</td>
<td>0.122</td>
<td>0.041</td>
<td>0.041</td>
<td>0.041</td>
<td>0.041</td>
</tr>
<tr>
<td>0.05</td>
<td>0.109</td>
<td>0.037</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
</tr>
<tr>
<td>1</td>
<td>0.186</td>
<td>0.010</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td>2</td>
<td>0.140</td>
<td>0.010</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td><strong>College</strong></td>
<td><strong>β</strong></td>
<td><strong>p</strong></td>
<td><strong>β</strong></td>
<td><strong>p</strong></td>
<td><strong>β</strong></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td><strong>β</strong></td>
<td><strong>p</strong></td>
<td><strong>β</strong></td>
<td><strong>p</strong></td>
<td><strong>β</strong></td>
</tr>
<tr>
<td><strong>Threshold</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
<td><strong>4 × 10^{-4}</strong></td>
</tr>
<tr>
<td>0.01</td>
<td>0.122</td>
<td>0.041</td>
<td>0.041</td>
<td>0.041</td>
<td>0.041</td>
</tr>
<tr>
<td>0.05</td>
<td>0.109</td>
<td>0.037</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
</tr>
<tr>
<td>1</td>
<td>0.186</td>
<td>0.010</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td>2</td>
<td>0.140</td>
<td>0.010</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
</tr>
</tbody>
</table>

**TABLE I.** The Standardized Regression Coefficients (β) for the Association Between the Polygenic Scores, Based on the Clumped Results for Years of Education (EduYears) and Completion of a College Degree (College), for Educational Achievement, School Performance, Attention Problems (AP) and ADHD Symptoms, Rated by Teacher and Mother.
FIG. 1. The proportion of variance explained ($R^2$) in Educational Achievement, School Performance and Attention Problems (AP)/ADHD symptoms by the polygenic scores based on the clumped results for years of education (EduYears) calculated for different p-value thresholds.
FIG. 2. The proportion of variance explained ($R^2$) in Educational Achievement, School Performance and Attention Problems (AP)/ADHD symptoms by polygenic scores based on the clumped results for completion of college degree (College) calculated for different p-value thresholds.
et al., 2010). This finding implies that at least some of the knowledge obtained with molecular genetic studies towards educational achievement can be used in research towards ADHD.

On the bases of the findings from the polygenic score analyses we cannot establish whether the observed effects of the genetic variants on both educational achievement and ADHD are direct or indirect. It could be that shared underlying biological mechanisms are the cause of the association between educational achievement and ADHD symptoms. When there is pleiotropy some of the genetic variants have a true direct biological influence on both phenotypes. For example, a number of molecular genetic studies demonstrated that there are some genetic variants on chromosome 6, 13 and 14 that have an effect on both reading disability and ADHD [Willcutt et al., 2002; Gayán et al., 2005]. The pleiotropic effect of a genetic variant can occur when a gene is involved in multiple biological pathways or the same biological pathway has different effects on the associated phenotypes [Solovieff et al., 2013]. For example, dysfunction in the dopaminergic pathway has been implicated in the development of ADHD and this pathway has also been associated with cognitive function [Nieoullon, 2002]. Alternatively, it may be that the genetic association appears because there are genetic variants influencing ADHD and, being genetically predisposed to ADHD makes it harder to concentrate at school, leading to lower educational achievement. Or, the other way around, children who have problems keeping up in school display, perhaps out of boredom and frustration, ADHD symptoms. The latter hypothesis seems to be refuted by findings from earlier studies that demonstrated that ADHD symptoms measured before children attended primary school are already associated with lower educational achievement years later [McGee et al., 1991; Fantuzzo et al., 2003].

The association between the polygenic scores and ADHD symptoms depended on the measurement instrument and rater. The effects were more strongly related, especially for the mother ratings, to the ADHD index of the CTRS-R/CPRS-R and not to the AP scale of the TRF/CBCL. One possible explanation is that the items of the AP scales are less school-oriented and include items less related to ADHD symptoms as described in the DSM than the ADHD index [Achenbach, 1991; Conners et al., 1998; APA, 2000].

The proportion of variance explained by the polygenic scores is small. According to a series of power analyses, this is as expected given the small effects of the individual genetic variants attributed to the complexity of the phenotypes [Dudbridge, 2013]. It is also consistent with findings in similar studies using polygenic scores to predict other complex phenotypes, for example, schizophrenia [International Schizophrenia Consortium, 2009] and intelligence [Davies et al., 2011]. Due to the highly polygenic nature of these phenotypes, the effects of the genetic variants are small and the standard errors of the estimates of the effect sizes are relatively large, limiting the predictive power of the polygenic score. Furthermore, the idea behind polygenic scores is that all genetic variants, also the ones that are not robustly associated with the phenotype, are included in the prediction, resulting in a majority of genetic variants without a true effect.

To conclude, the findings of shared genetic variants between educational achievement and ADHD may lead to new possibilities in the research towards ADHD. Larger sample sizes to identify genetic variants influencing ADHD might be within reach, since data on educational achievement are already available at a larger scale for participants with genotype data as well as easier to collect than ADHD data.

ACKNOWLEDGEMENTS
We are grateful to the twin families and the teachers for their participation. This project is part of the research program “Innovative learning materials and methods” funded by the Netherlands Initiative Brain and Cognition, a part of the Organization for Scientific Research (NWO) under grant 056-31-001. This research was supported by “Spinozapremie” (NWO/SP1 56-464-14192); “Twin-family database for behavior genetics and genomics studies” (NWO 480-04-004); Genetics of Mental Illness: European Research Council (ERC-230374); Genetic influences on stability and change in psychopathology from childhood to young adulthood (NWO/ZonMW 91210020); Genomics of Developmental Trajectories in Twins (ARRA RC2 2MH08995).

APPENDIX
The following people who are not listed as co-authors on this manuscript contributed to the original GWAS meta-analysis on educational attainment [Rietveld et al., 2013], on which the present paper is based. Data access has been granted under section 4 of the Data Sharing Agreement of the Social Science Genetic Association Consortium (SSGAC). The views presented in the present paper may not reflect the opinions of the individuals listed below. The SSGAC is grateful to the authors of [Rietveld et al., 2013] for providing the meta-analysis data. We thank: Abdel Abdellaoui, Arpna Agrawal, Eva Albrecht, Behrooz Z. Alizadeh, Jüri Allik, Najaf Amin, John R. Attia, Stefania Bandinelli, John Barnard, François Bastardot, Sebastian E. Baumeister, Jonathan Beauchamp, Daniel J. Benjamin, Kelly S. Benke, David A. Bennett, Klaus Berger, Lawrence F. Bielak, Laura J. Bierut, Jeffrey A. Boatman, Dorret I. Boomsma, Patricia A. Boyle, Ute Bültmann, Harry Campbell, David Cesarini, Christopher F. Chabris, Lynn Cherka, Mina K. Chung, Dalton Conley, Francesco Cucca, George Davey-Smith, Gail Davies, Mariza de Andrade, Philip L. De Jager, Christiana De Leeuw, Jan-Emmanuel De Neve, Jan I. Deary, George V. Dedoussis, Panos Deloukas, Jaime Derringer, Maria Dimitriou, Gudny Eiriksdottir, Niina Eklund, Martin F. Eldred, Johan G. Eriksson, Tönö Esko, Daniel S. Evans, David M. Evans, Jessica D. Faul, Rudolf Fehrmann, Luigi Ferrucci, Krista Fischer, Lude Franke, Melissa E. Garcia, Christian Gieger, Häkon J. Gjessing, Patrick J.F. Groenen, Henrik Grönborg, Vilmundur Gudnason, Sara Hägg, Per Hall, Jennifer R. Harris, Juliette M. Harris, Tamara B. Harris, Nicholas D. Hastie, Caroline Hayward, Andrew C. Heath, Dena G. Hernandez, Wolfgang Hoffmann, Adriaan Hofman, Albert Hofman, Rolf Holle, Elizabeth G. Holliday, Christina Holzapfel, Jouke-Jan Hotenga, William G. Iacono, Carla A. Ibrahim-Verbaas, Thomas Illig, Erik Ingelsson, Bo Jacobsson, Marjo-Riitta Jarvelin, Min A. Jhun, Magnus Johannesson, Peter K. Joshi, Astanand Jugessur, Marika Kaakinen, Mika Kähönen, Stavroula Kanoni, Jaakko Kaprio, Sharon L.R. Kardia, Juha Karjalainen, Robert M. Kirkpatrick, Philipp D. Koellinger, Ivana Kolcic, Matthew Kowgier, Kati Kristiansson, Robert F. Krueger, ZoltanKutalik, Jari Lahti, David


